Large Scale Density Estimation of Blue and Fin Whales (LSD)

Jennifer L. Miksis-Olds
Applied Research Laboratory
The Pennsylvania State University
PO Box 30 Mailstop 3510D
State College, PA 16804
phone: (814) 865-9318 fax: (814) 863-8783 email: jlm91@psu.edu

Award Number: N00014-14-1-0397

Len Thomas & Danielle Harris
Centre for Research into Ecological and Environmental Modelling
University of St Andrews
The Observatory, Buchanan Gardens
St Andrews Fife, KY16 9LZ, Scotland, UK
phone: (0)1334-461801 fax: (0)1334-461800 email: len.thomas@st-andrews.ac.uk

Award Number: N00014-14-1-0394

LONG-TERM GOALS

Effective management and mitigation of marine mammals in response to potentially negative interactions with human activity requires knowledge of how many animals are present in an area during a specific time period. Many marine mammal species are relatively hard to sight, making standard visual methods of density estimation difficult and expensive to implement; however many of these same species produce vocalizations that are relatively easy to hear, making density estimation from passive acoustic monitoring data an attractive, cost-effective alternative. A particularly efficient passive acoustic monitoring design is a "sparse array", where sensors are distributed evenly over a large area of interest – however a consequence of this design is that each vocalization cannot be heard at multiple sensor locations, restricting the choice of methods that can be used to estimate density. Nevertheless, sparse array methods have been developed and demonstrated (Marques et al., 2011, Küsel et al., 2011; Harris, 2012; Harris et al., 2013). While these studies represent an important step forward in making the methods more generally applicable at reasonable cost, they have some drawbacks: they either are only applicable to small local ocean areas, or they require unrealistic assumptions about animal distribution around the sensors, or both. The goal of this research is to develop and implement a new method for estimating blue and fin whale density that is effective over large spatial scales and is designed to cope with spatial variation in animal density utilizing sparse array data from the Comprehensive Nuclear Test Ban Treaty Organization International Monitoring System (CTBTO IMS) and Ocean Bottom Seismometers (OBSs).

OBJECTIVES

This effort will first develop and implement a density estimation methodology for quantifying blue and fin whale abundance from passive acoustic data recorded on sparse hydrophone arrays in the Equatorial Pacific Ocean at Wake Island. It builds on previous work with sparse arrays of OBSs. Density estimation methods developed in the Pacific Ocean at Wake Island will then be applied to the same species in the Indian Ocean at the CTBTO location at Diego Garcia.

Specific objectives are as follows.

1. Develop and implement methods for estimating detection probability of vocalizations based on bearing and source level data from sparse array elements.

2. Validate using OBS data, where additional independent information on detectability is available.

3. Use all available and relevant data to develop multipliers for converting calls-per-unit-area to blue and fin whale density – i.e., estimates of average call rate.

4. Estimate the regional density and spatial distribution of blue and fin whales in the Equatorial Pacific Ocean, using CTBTO data from Wake Island.

5. Estimate regional density and spatial distribution of blue and fin whales in the Indian Ocean, using CTBTO data from Diego Garcia.

APPROACH

Researchers at the Applied Research Laboratory at Penn State (ARL Penn State) are working collaboratively with the Centre for Research into Ecological and Environmental Modeling (CREEM) at the University of St. Andrews. The St. Andrews team provides expertise in density estimation techniques from passive acoustic datasets, while collaborators at ARL Penn State provide the long-term data series and expertise in marine mammal biology, acoustic processing, ambient sound, and sound propagation. This project leverages multiple research products from previous and current funding from ONR, Navy Living Marine Resources (LMR) Program, NOAA, JIP, and the UK Defense Science and Technology Laboratory (DSTL). Low frequency (1-120 Hz), continuous data recorded by the CTBTO IMS for over or close to a decade at Diego Garcia (2002-present: Indian Ocean), and Wake Island (2007-present: Equatorial Pacific Ocean) have been acquired under a current ONR YIP Award N000141110619 to Miksis-Olds (ARL PSU). A near real-time portal has been opened between ARL PSU and the AFTAC/US NDC (Air Force Tactical Applications Center/ US National Data Center) to continue to download data from these two locations. The density estimation method development builds on the work of Danielle Harris (PhD work funded by UK DSTL; Cheap DECAF project funded by ONR N00014-11-1-0615) and Len Thomas (DECAF project, funded by NOAA and JIP through NOPP).

The CTBTO IMS instrument configuration of hydrophone triads suspended in the deep sound channel allows for call bearing and, in some cases where the vocalizing animal is close, localization (Harris 2012; Samaran et al., 2010). This, together with received level, potentially allows the distribution of animals to be estimated without requiring randomly placed multiple instruments. It is anticipated that bearings and received levels of a large number of calls can be estimated. We plan to use these data, coupled with estimates of call source levels and sound propagation models in the study area, to

estimate the distribution and density of calling whales in the monitored area. To do this, we will use the bearing, source level and transmission loss estimates to estimate the location of each call and range over which calls can be detected (together with estimates of uncertainty on these quantities). A detailed detector characterization will give probability of detection as a function of signal-to-noise ratio (SNR), and hence we can estimate probability of detection for each received call. Spatio-temporal variability in the efficiency of the automatic detector will also be considered. Call "abundance" at the location of each call can then be estimated with a Horvitz-Thompson-like estimator, where each detected call is scaled by its associated probability of detection to account for undetected calls also produced at that location (Borchers et al., 2002; Thompson 2012). The resulting estimates will be smoothed in space with a Generalized Additive Model (GAM) to give an estimated density surface (Wood, 2006). Taken together, this represents a novel approach to density estimation that has wide applicability.

Density estimation from passive acoustic recorders relies heavily on the detection of vocalizations above the noise and knowledge of the acoustic coverage (or active acoustic space) of each passive acoustic sensor. Estimation of the range of acoustic detection is a function of signal source level, SNR of detection, and sound propagation. Sound propagation characteristics and ambient noise dynamics are site specific and highly time dependent, so an acoustic propagation model that incorporates the changing acoustic and oceanographic conditions will be applied to calculate the acoustic coverage over time for each sensor. Noise level is likely the most variable factor affecting the range of acoustic detection. Sound levels at Wake Island over the past 5 years show frequency-dependent seasonal patterns (Miksis-Olds et al., 2014), so a seasonal component will be included in the optimal acoustic coverage model. SNR detection thresholds will be established at each site for both a north and south array element. SNR detection threshold will be assessed on a subset of calls each year and monitored over the duration of the dataset to assess any long-term changes. There is evidence that tonal blue whale calls are decreasing in frequency over time (McDonald et al., 2009), which is why it will be necessary to verify SNR detection thresholds and adjust detectors as needed.

In addition to understanding the time-varying environmental components influencing call detection, use of the most appropriate source levels is critical to computing accurate detection ranges and final density estimations. Localized calls (from nearby animals) on a given CTBTO array will provide a distribution of regional source level estimates. This will be preferable to source level estimates taken from the literature. The proposed density estimation method is also highly dependent on call rate inputs, which are used in the development of species specific multipliers for converting the number of detected calls to the estimated number of animals. Blue and fin whale call rates are best estimated from tagged animals, and DTag (digital acoustic tag) data are available for blue and fin whales through ongoing ONR projects, where we are currently communicating with the project PIs to acquire realistic call rate information.

Quantifying uncertainty in estimates is as important as obtaining the estimates themselves. Our inputs to the acoustic modeling will be a distribution on source level, and will include quantification of measurement error in bearing. Uncertainty in these inputs will be cascaded through the acoustic modeling, and combined with variance estimates for detector performance and call rates to provide a robust estimate of uncertainty in density. An example of this kind of uncertainty propagation is given by Harris (2012, Chapter 6).

The use of bearing data is a new density estimation methodology, and we will use OBS array data in a pilot study. An array of 24 instruments was deployed off the coast of Portugal for 12 months in

2007/2008. Each OBS has a sampling rate of 100 Hz and many fin whale calls have been detected (Harris, 2012). Both range and bearing to each call can be estimated using the OBS array (Harris et al., 2013), providing an ideal dataset with which to compare the new method with an existing robust density estimation method. Using this array, density results obtained using bearing data can be directly compared with density results obtained using standard distance sampling.

WORK COMPLETED

A project progress meeting took place at Penn State July 6-10, 2015. PI Miksis-Olds (ARL), Post-doc Harris (CREEM), and graduate student Julia Vernon (ARL) participated. This meeting focused on the data analysis and method development that took place over the past year relating to the pilot study. An outline for and timeline for two papers related to the pilot study was developed: 1) methods paper, and 2) short-range validation paper comparing density estimation results from Wake Island CTBTO data and OBS data. The meeting concluded with a discussion on tasks and responsibilities related to moving forward with the second phase of study in assessing long term density estimation patterns at Wake Island over the past eight years.

The proposed method for using bearing and SNR data to estimate abundance and density has been developed using R, an open-source statistical software package (R Core Team, 2014). Both a simulation and analysis tool have been created. The simulation tool allows users to run simulations specific to their study site, study species, and detection process. This allows an assessment of the size of the monitored area, given the signal of interest's source level, local transmission loss properties, and the efficiency of the automatic detector. The simulation tool also allows users to assess the level of bias that may occur at the data analysis stage and at what monitoring range the bias is minimized. Simulations can be developed for different distributions of animals around the instrument. The analysis tool uses the same method implemented in the simulation code, but allows users to input their collected survey data.

The team has been working with a 3 month time period from December 2007-February 2008 for the pilot study at Wake Island in the Equatorial Pacific Ocean. This time period provides complete overlap between the CTBTO IMS data and the OBS data. Fin whales were identified as the target species for the pilot study. Two different types of automatic detectors were considered and assessed for application to the pilot and long term studies: matched filter and spectrogram correlation. The matched filter detector cross-correlates the time waveform of the desired signal (a fin whale call) with the time waveform of the dataset. The spectrogram correlation method involves cross-correlating the spectrogram of the dataset with a synthetic kernel. The kernel is a template that indicates the time and frequency endpoints of the desired call. Results from automatic detectors were compared with manually detected calls over the duration of the pilot study. In the case that an automatic detector indicated a call that was not detected manually, the detected signal was marked as a false positive detection. The false positive rate is then the number of false positive detections divided by the total number of automatic detections. In the case that a call was detected manually but not automatically, the call was marked as a missed call. The proportion of missed calls to total number of manually detected calls was then determined. ROC curves were generated by varying the sensitivity of the detectors and determining the corresponding false positive rates and percentage of calls detected. Once an optimal detector was developed, the relationship between call signal to noise ratio and probability of detection was determined by fitting a GAM to SNR data from both detected and undetected calls.

In preparation for the analysis of the pilot study data, bearings to detected fin whale calls in the pilot study were calculated through use of time difference of arrival (TDOA) of received signals. TDOA between each pair of hydrophones (N1 and N2, N2 and N3, N3 and N1) were found by cross-correlating the received signals. For some calls distortion in the waveform prevented cross-correlation and time delays were found manually. Using the known distances between receivers and the sound speed, an estimated bearing was calculated for each pair of hydrophones. The median was then selected from the three bearing pairs.

Source level calculations were also performed on vocalizations manually detected from spectrograms of the CTBTO data. Received levels were calculated, and noise level at the time of the call was also obtained from the CTBO time series. Transmission loss (TL) was determined using a site- and season-specific OASIS Peregrine parabolic equation model. The model incorporates the location of the sensor in the deep sound channel, the bathymetry of the area and the local sound speed profiles. TL was modeled for 360 bearings with a 1 degree resolution. TL values between the sensor and source were found for individual vocalizations using ranges and bearings calculated through hyperbolic localization. The exact depths of the sources were unknown but assumed to be within the upper 300 m of the water column. Source levels were then calculated using the passive sonar equation.

RESULTS

Simulations have been run for both fin and blue whales using transmission loss data from Wake Island and Diego Garcia. Examples of simulation results are presented in Figures 1 – 3, which are based on a blue whale study at Diego Garcia. The simulation was run 100 times and the results suggested that the median bias in estimated abundance or density expected in an analysis of data collected from a similar survey scenario was -7.2%. The simulation results also gave a median optimal monitoring range of 820 km; bias was reduced at this range to a median level of 0.1%. Further runs, with higher numbers of replications and under more diverse conditions are planned.

Fin whale vocalizations detected during the pilot study resulted in 1484 detections ranging in received level from approximately 100-132 dB re 1 μPa (Figure 4). The spectrogram correlation detector was determined to be the most effective detector for fin whales at this location. A spectrogram correlation detector with a 10 % false positive rate was identified as the optimal detector for fin whales at the Wake Island location. A 10% false positive rate maximized the percentage of calls detected while minimizing the percentage of false positives included (Figure 5).

Modelling detection probability as a function of SNR predicted that detection probability was less than 0.1 at a SNR of 5 and any signal with a SNR of over 15 was likely to be detected with certainty (Figure 6).

Source levels were determined from whales detected in close proximity to the array during the pilot study period as well as for time periods extending outside the pilot study. A mean SL of 188.9 dB re 1 μPa (+/- 2.6 dB re 1 μPa) and median level of 189.2 dB re 1 μPa was determined from 170 identified calls (Figure 7).

Bearings calculated during the pilot study clearly showed that the fin whale distribution was not uniform around the Wake Island North array (Figure 8). As analysis transitions to analysis of the entire dataset, a subset of bearings will be calculated each year to examine interannual differences.

IMPACT/APPLICATIONS

Acoustic monitoring for the presence of marine life is an ongoing Navy need in meeting regulatory requirements, and offers a low cost alternative to visual surveys. The density estimation method developed here for the targeted low frequency vocalizations of blue and fin whales will be directly applicable to other species and frequency ranges using sparse arrays of fixed or remotely deployed PAM systems. Outputs will be of direct relevance to Navy risk assessment models.

TRANSITIONS

To be determined as this project unfolds.

RELATED PROJECTS

The propagation modeling included in this study in collaboration with Kevin Heaney (OASIS) is directly related to ONR Ocean Acoustics Award N00014-14-C-0172 to Kevin Heaney titled "Deep Water Acoustics".

The current project is also directly related to and follows on to ONR Award N000141110619 to Jennifer Miksis-Olds titled "Ocean Basin Impact of Ambient Noise on Marine Mammal Detectability, Distribution, and Acoustic Communication". Patterns and trends of ocean sound observed that study will be directly applicable to the estimation of signal detection range in this study.

The density estimation method development builds on the work of Danielle Harris (PhD work funded by UK DSTL; Cheap DECAF project funded by ONR N00014-11-1-0615) and Len Thomas (DECAF project, funded by NOAA and JIP through NOPP).

Result from tagging studies under ONR Award N00014-14-1-0414 "Behavioral context of blue and fin whale calling for density estimation" to Ana Širović will better inform the species specific multipliers for converting number of vocal detections into number of animals by providing information on source level and call rates.

REFERENCES

Borchers, DL, Buckland, ST and Zucchini, W (2002). *Estimating Animal Abundance.* Springer, New York.

Harris, D, Matias, L, Thomas, L, Harwood, J and Geissler, WF (2013). Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic. Journal of the Acoustical Society of America 134, 3522-3535.

Harris, D (2012). Estimating whale abundance using sparse hydrophone arrays. PhD Thesis: University of St. Andrews.

Küsel, ET, Mellinger, DK, Thomas, L, Marques, TA, Moretti, D, and Ward, J (2011). Cetacean population density estimation from single fixed sensors using passive acoustics. The Journal of the Acoustical Society of America 129, 3610-3622.

Marques, TA, Munger, L, Thomas, L, Wiggins, S and Hildebrand, JA (2011). Estimating North Pacific right whale (*Eubalaena japonica*) density using passive acoustic cue counting. Endangered Species Research 13, 163-172.

McDonald, MA, Hildebrand, JA, and Mesnick, S (2009). Worldwide decline in tonal frequencies of blue whale songs. Endangered Species Research 9, 13–21.

Miksis-Olds, JL, Vernon, JA and Heaney, K (2014). Applying the dynamic soundscape to estimates of signal detection. Proceedings of the 2014 Underwater Acoustics International Conference and Exhibition, Rhodes, Greece, June 22-27, 2014.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Samaran, F, Adam, O, and Guinet, C (2010). Detection range modelling of blue whale calls in the Southwestern Indian Ocean. Applied Acoustics 71: 1099-1106.

Širović, A, Hildebrand, JA, Wiggins, SM (2007). Blue and fin whale call source levels and propagation range in the Southern Ocean, J. Acoustical Soc. of Am. 122: 1208-1215.

Thompson, SK (2012). *Sampling,* 3rd Edition. Wiley.

Wood, SN (2006). *Generalised Additive Models: An Introduction with R.* Chapman & Hall, Boca Raton, FL.

PRESENTATIONS

Harris, D, Thomas, L, Miksis-Olds, JL, Vernon, JA (2015). Large scale density estimation of blue and fin whales: combined distribution and density estimates using bearing data. 7th International Workshop on Detection, Classification and Localization of Marine Mammals Using Passive Acoustics. La Jolla, California. July 13-26.

Vernon, JA (2015). Automatic detection and bearing calculation of vocalizing marine mammals in relation to passive acoustic density estimation. Imagining the Future of Ocean Science Symposium. Center for Marine Science & Technology, Penn State. 14 July, 2015.

Vernon, JA, Miksis-Olds, JL, Harris, D (2015). Analysis of bearings of vocalizing marine mammals in relation to passive acoustic density estimation. 170th Meeting of the Acoustical Society of America. Jacksonville, FL, November 2-6.

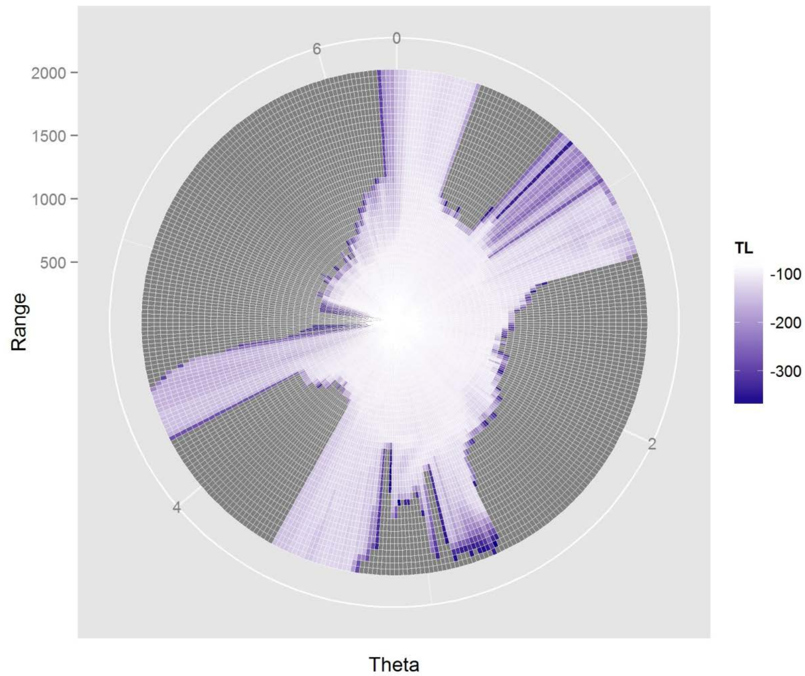

Figure 1. Modelled transmission loss data from Diego Garcia. The instrument is located in the middle of the plot. The grayed areas depict transmission loss levels that were considerd infinite by the propagation model. Therefore, blue whale calls produced in these areas would be masked and unable to be received on the hydrophone.

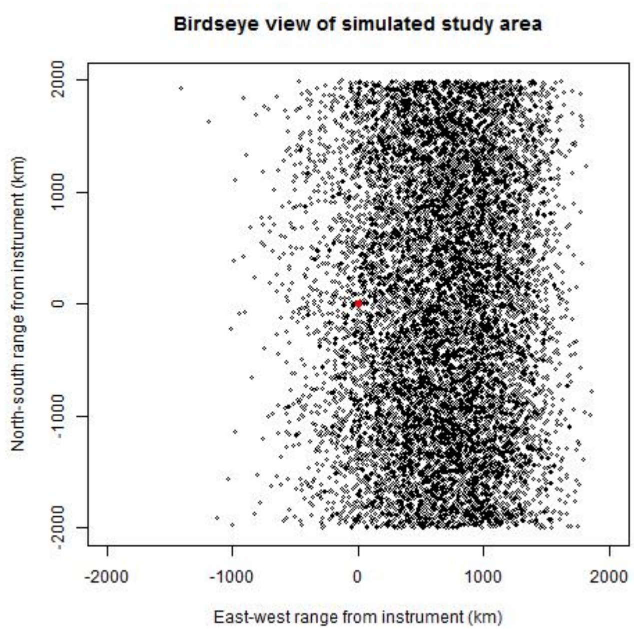

Figure 2. Example of a simulated non-uniform animal distribution in a 2000 km x 2000 km area. The instrument location is depicted by the red dot at location (0,0).

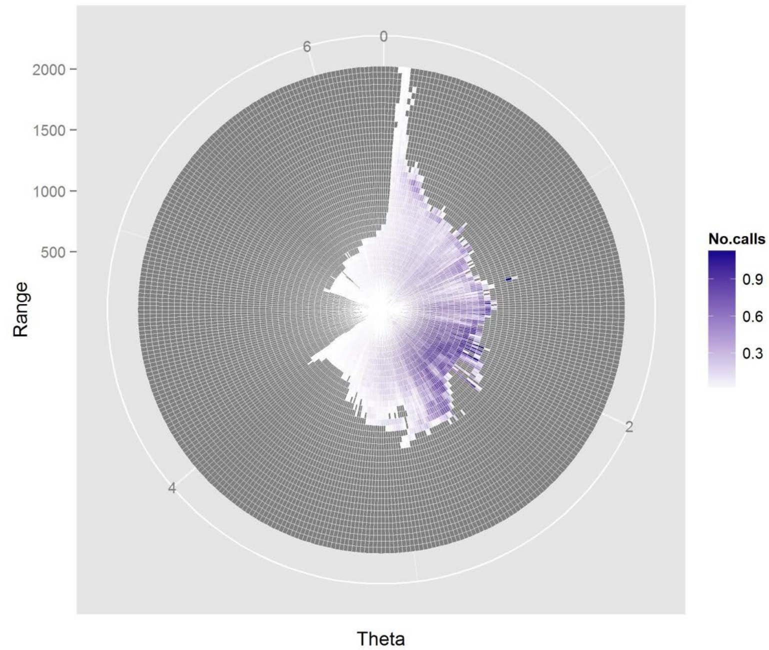

Figure 3. Predicted call abundance around the instrument based on analysing bearing and SNR data from detected calls in the simulated population. Note that the masked area has increased in comparison to Figure 1 - calls in further areas are considered masked once data about the efficiency of the automatic detector are incorporated into the analysis.

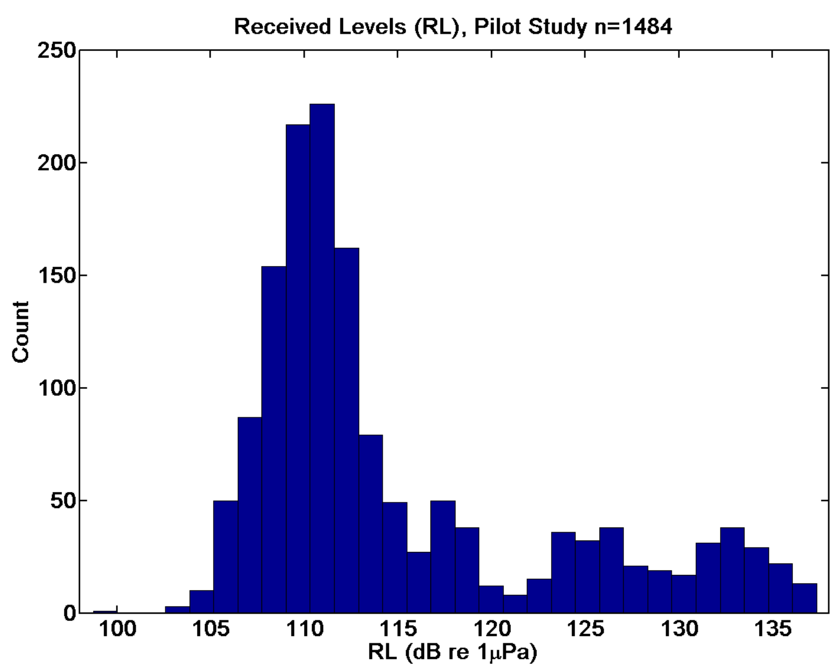

Figure 4. Received levels of fin whale calls manually detected in the pilot study (n = 1484).

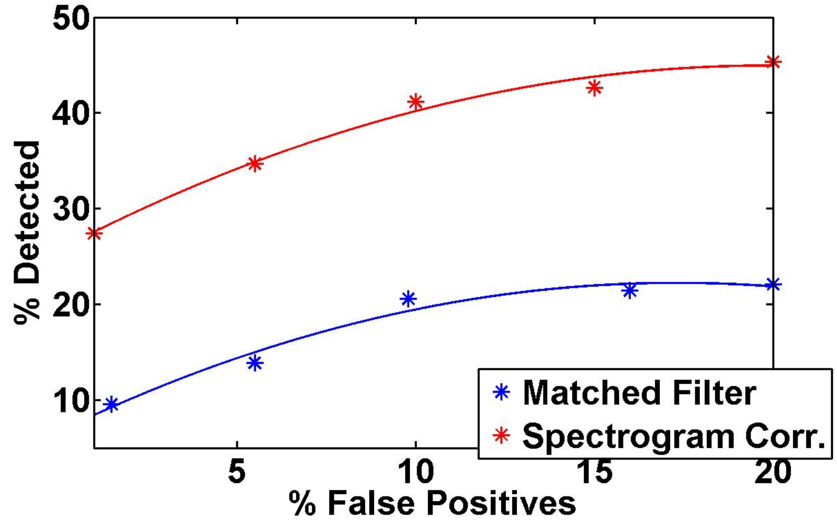

Figure 5.   Receiving Operator Characteristic (ROC) curves for two methods of automatic detection of fin whale calls. The ROC curves were developed using the calls detected in Figure 4.

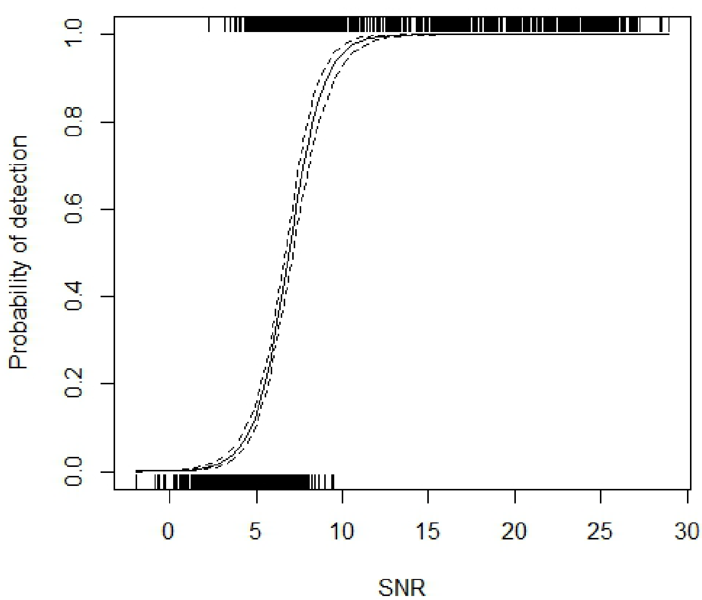

Figure 6.   Detector characterization curve linking SNR to detection probability. The curve is estimated using a Generalized Additive Model fitted to SNR data of both automatically detected, and undetected, calls. The dotted lines around the curve depict the 95% confidence limits. The vertical lines above and below the curve are a rug plot, showing the SNR values of detected calls (lines above the curve) and undetected calls (lines below the curve)

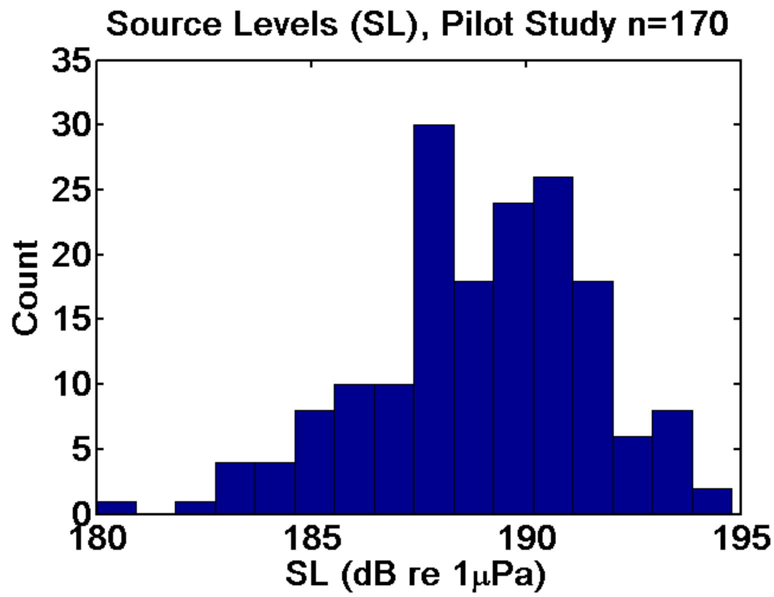

Figure 7. Source levels of fin whale calls measured from the pilot study dataset (n = 170).

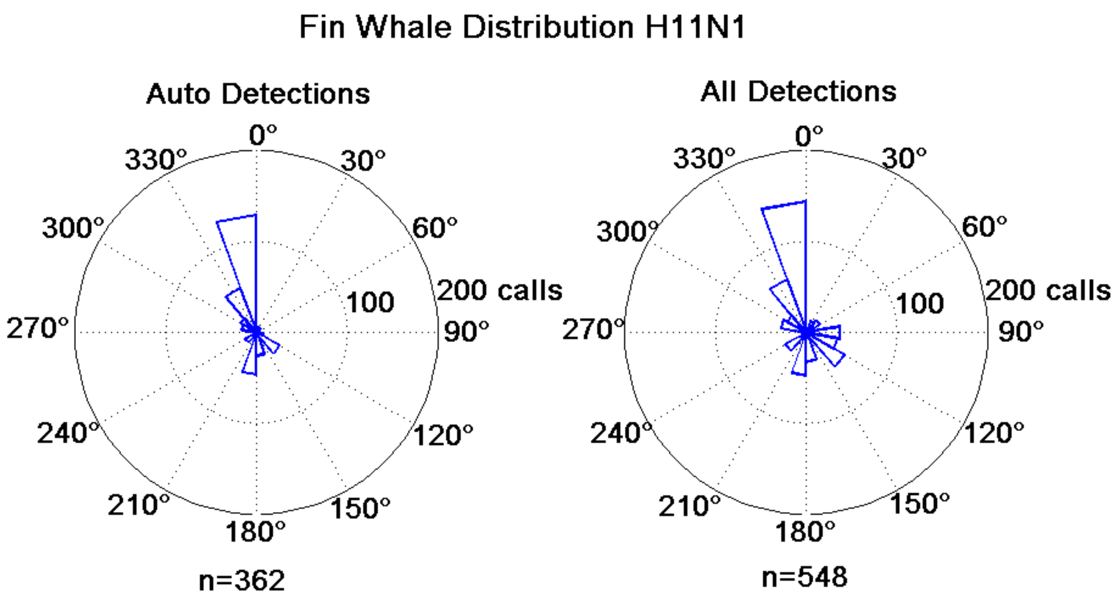

Figure 8. Estimated bearings of fin whale calls from the pilot study dataset (n = 548). Bearing data are shown for (1) calls detected by the automatic detector and (2) all calls (whether detected or undetected by the automatic detector).